TRANS-PACIFIC DEMONSTRATIONS (TPD)
LESSONS LEARNED

DRAFT

Richard desJardins, NASA Research and Education Network
Ray Gilstrap, NASA Research and Education Network
Paul Lang, NASA Goddard Space Flight Center
Naoto Kadowaki, Communications Research Laboratory
Eddie Hsu, NASA Jet Propulsion Laboratory
Haruyuki Tatsumi, Sapporo Medical University
Mike Gill, National Library of Medicine

Version 1.0

November 2000

http://www.nren.nasa.gov/tpd

CONTENTS

INTRODUCTION—TRANS-PACIFIC DEMONSTRATIONS

The Trans-Pacific Demonstrations (TPD) were a coordinated set of prototyping developments, measurements, tests, and live demonstrations carried out from May to July 2000 by US, Japanese and Canadian organizations interested in supporting scientific activities over networks that include long delay satellite links. This document reports on the engineering, scientific, and management lessons learned during the preparation and conduct of these demonstrations. For more information, visit the TPD web page at

http://www.nren.nasa.gov/tpd

1 VISIBLE HUMAN

One demonstration involved interactive access to the US National Library of Medicine (NLM) Visible Human anatomical digital library. This demonstration involved the use of sophisticated distributed processing software to compose the query results for an interactive anatomical slice request, and included a double satellite hop with an end-to-end round-trip distance of over 300,000 kilometers between Bethesda, USA and Sapporo, Japan. The main issues for this demonstration centered around the long delay caused by the two-satellite hop over Intelsat and N-Star, and the use of a distributed file system to access the data.

2 REMOTE ASTRONOMY

Another demonstration involved Japanese high school students interactively requesting and receiving astronomical images in real time from telescopes located at Mount Wilson in California. The main issues for this demonstration included the effect of long latency on student interactivity and the use of videoconferencing and audioconferencing technologies for real time operations.

3 ENGINEERING

TPD engineering issues included network configuration and operations as well as the end-to-end protocols and software needed to support the end-system needs of the demonstration applications.

4 MANAGEMENT

TPD management issues included programmatic and budgeting issues as well as the administrative and operational coordination of the many organizations and activities involved in TPD.

5 ORGANIZATION OF THIS DOCUMENT

Necessarily, organization of a lessons learned document is problematical, especially for TPD, because lessons learned cut across any taxonomy used to organize the activities. This lessons learned document is a compilation of inputs collected from the main organizations involved in TPD.

6 ACKNOWLEDGEMENTS

Sincere thanks are due to Intelsat, BCnet, Teleglobe, Canarie, and APAN/TransPAC, as well as the participants from the Japanese and US organizations noted in the remainder of this report. TPD was difficult and challenging, and as it turns out, would have been a complete failure without the dedicated efforts of the team of 50 or so professionals from these organizations. Many of the TPD team were extremely giving of their personal time and efforts, which in some cases went far beyond their organization's commitments. We thank all members of the TPD team, individually and collectively, for bringing TPD to a successful conclusion with the publication of this lessons learned report.

NETWORK ENGINEERING —TRANS-PACIFIC DEMONSTRATIONS

Ray Gilstrap, NASA Research and Education Network

1 BACKGROUND

This report presents the results of network-level tests conducted in support of the Trans-Pacific Demonstrations (TPD). For the Visible Human (VH) demonstration, the TPD network connected an application server at the National Library of Medicine (NLM) with clients at the Sapporo Medical University (SMU). For the Remote Astronomy (RA) application, a telescope at Mt. Wilson Observatory and an application server at NASA Ames Research Center (ARC) were connected to clients at the Communications Research Laboratory (CRL), the University of Maryland – College Park, and NASA Jet Propulsion Laboratory.

2 NETWORK ARCHITECTURE

The attached figures (see http://www.nren.nasa.gov/tpd) show an overview of the network. Key end sites included:

- National Library of Medicine (Bethesda, Maryland) -- VH server site
- Sapporo Medical University (Sapporo, Japan) -- VH client site
- NASA Ames Research Center (Moffett Field, California) -- RA server site
- Mount Wilson Observatory (California) -- RA telescope site
- Communications Research Laboratory (Tokyo, Japan) -- RA client site
- NASA Jet Propulsion Laboratory (Pasadena, California) -- RA client site
- University of Maryland (College Park, Maryland) – RA client site

Key points in the network infrastructure included:

- NASA Goddard Space Flight Center (Greenbelt, Maryland) – NREN/HPCC Point-of-Presence
- STAR TAP (Chicago, Illinois) -- North America Exchange Point
- Tokyo XP (Tokyo, Japan) -- Japan Exchange Point for APAN
- Lake Cowichan, British Columbia – Satellite Earth Station
- Kashima Space Research Center (Tokyo, Japan) – Satellite Earth Station
- BC GigaPOP (Vancouver, British Columbia) – Canada Exchange Point

Primary wide-area transit service was provided by:

- NREN (U.S.)
- Teleglobe (Canada)
- AT&T Canada (formerly Metronet)
- CA*net3 (Canada)
- APAN/TransPAC (U.S.-Japan)

- IMnet (Japan)

Phase 1 of the demonstration (functional testing of applications) was conducted via the APAN/TransPAC link. Phase 2 (operational phase) was conducted over the Intelsat satellite. Additionally, the Visible Human application transited N-Star, a second intra-Japan satellite.

Key features of the network architecture included:

- ***Parallel paths*** –The satellite and terrestrial paths were both available simultaneously. Routing information for the U.S.-side servers was advertised to both TransPAC and CA*net3 by the NREN Chicago router. Likewise, this router received routing information for the Japan-side clients from both CA*net3 and TransPAC. The CA*net3 routes were preferred. When the satellite became temporarily unavailable, traffic automatically redirected to the TransPAC path.
- ***SkyX translation*****—**For the Visible Human application, the SkyX TCP connection enhancement gateway was used. SkyX replaces TCP with a proprietary protocol that is optimized for reliable transmission over high bandwidth-delay links such as satellites. SkyX gateways were placed at GSFC and at SMU so that delays over both the Intelsat and N-Star satellites could be compensated for. More details of SkyX-related issues are available in a separate report.

3 NETWORK TESTS

Once the initial connectivity was established, tests were conducted to determine reachability, round trip times, loss rates, hop counts, throughput, and path symmetry along various segments of the network.

The primary test paths for all of the tests included the following:

NLM-SMU*

ARC-MWO

ARC-CRL*

ARC-JPL

ARC-UMD

Paths marked with an asterisk were tested over both the APAN/TransPAC link and the satellite link.

3.1 Reachability/Round Trip Time/Jitter/Packet Loss

These characteristics were measured by running a series of pings from a UNIX workstation at each site to workstations at all of the other sites. The ping tests were automatically run on each host every four hours, and the results were logged to a central server and made available on the Web. The aggregate of these data can be used to build a profile of round trip times across different segments of the network. The delay experienced over the Trans-Pacific link (via both APAN and Intelsat) was of particular interest, as well as the delay over the N-Star link.

3.2 Hop Counts/Congestion/Path Symmetry

These characteristics were measured by running a series of traceroutes between the test workstations, also at four-hour intervals. These data yield some information about route stability symmetry and potential congestion points.

3.3 Throughput

TCP and UDP throughput was measured using iperf. NFS, the distributed filesystem protocol used by the VH application, was run over TCP connections. As a result, the large bandwidth-delay product of the link was a major concern because the VH application ran on an operating system whose TCP stack could not be adequately tuned to provide maximum throughput; hence the need for the SkyX gateways. The RA application used the AFS distributed filesystem over UDP and was not subject to the same concerns.

3.4 Results

The following table provides a summary of the measurements:

Path	Via	Throughput (Mb/s TCP/UDP)	RTT (ms)	Jitter (ms)	# Hops	Loss Rate (%)	Sym-metric ?
NLM-SMU	TransPAC	1.2/1.3	192		17	0	Y
	Intelsat	15.3/32.1	1131		15	0	Y
SMU-NLM	TransPAC	1.2/1.3	192		16	0	
	Intelsat	11.9/32.0	1131				Y
ARC-CRL	TransPAC	0.26/9.0				0.1	Y
	Intelsat	0.27/10.0				0.5	Y
CRL-ARC	TransPAC	1.9/9.1					
	Intelsat	0.76/11.1					Y
ARC-JPL	NREN	NA/10.0*			7	0	
JPL-ARC	NREN	NA/NA	NA	NA	NA	NA	
ARC-UMD	NREN	3.9/7.7*	58		7	0	Y
UMD-ARC	NREN	6.2/NA	58		7	0	
MWO-ARC	NISN	1.2/1.3	19			0	Y
ARC-MWO	NISN	1.2/1.3	19			0	

The TCP measurements were taken with a window size of 64KB, which represents the maximum window size available on machines that do not support extended window sizes. Additional TCP throughput measurements were taken using a widow size of 8KB, the default on many operating systems. The results were slightly lower than the above.

The TCP measurements between NLM and SMU over Intelsat were taken with SkyX processing active. Without SkyX, the TCP results were 200-800 Kbps in each direction.

Values marked with an asterisk are estimates obtained using the ***bing*** tool which provides a rough estimate of throughput based on round trip times of different-sized ICMP packets. Bing is useful in situations where an iperf-type receiving daemon is not available on the remote end.

Of particular interest is the poor observed throughput in the TCP tests over the satellite link. The prevailing theory was that the low TCP throughput was caused by ATM cell drops, which resulted in lost TCP segments that triggered the TCP congestion control mechanisms. The cell drops were further posited to be caused by non-conformance to the rate limiting parameters set on the PVCs.

3.5 Additional Laboratory Tests

To test this theory, an experiment was set up in the NREN lab in which identical rate limiting was set up on the PVC connecting the two testbed routers. The parameters used were as follows:

Sustained Cell Rate (SCR)	52104 cps (20 Mbps)
Peak Cell Rate (PCR)	104208 cps (40 Mbps)
Maximum Burst Rate (MBR)	250 milliseconds
Cell Delay Variation Tolerance (CDVT)	250 microseconds

The MBR is the length of time in any given 1-second period that cells may be transmitted at the peak rate. CDVT is the maximum tolerable jitter. Cells that do not conform to these parameters are discarded.

FTP and iperf tests were run between ciocc2 and hang10. The results were similar to what was observed on the production circuits. The ATM cell counters showed a relatively high number of cell drops on the rate-limited PVC (~1200 for a 2.5 MB file). A log of tcpdump output showed a number of duplicate ACKs from the receiver back to the sender as TCP segments were lost. The duplicate ACKs caused the sender to repeatedly throttle its transmission rate, resulting in TCP throughputs in the range of 200-600 Kbps. By contrast, UDP throughput over the same PVC was consistently around 32 Mbps.

The results were repeated with different MBR values. As expected, there was a direct relationship between MBR and reliability. However, even when the MBR was set equal to 1 second (meaning that traffic could be continuously sent at the peak rate), reliability was observed also to depend on the CDVT value. A high number of cell drops occurred with the default CDVT value of 250 us. When this was increased to 1000 us, the cell drop count sharply diminished.

This suggests that the routers have difficulty generating a cell stream that conforms to the smaller CDVT value (i.e., with low jitter). The maximum reliability was observed with the MBR set to 1 second and the CDVT set to at least 1000 microseconds.

4 LESSONS LEARNED

4.1 Technical

The technical lessons learned from TPD illustrate the importance of the configuration of the network infrastructure at all layers, from physical to application. During the course of the project, performance concerns included:

- ***Propagation delay*** - This physical layer characteristic was well understood at the outset to have performance implications for the interactive sessions.
- ***ATM PVC parameters*** - The results of the performance tests on the TPD network, as well as the additional tests in the NREN lab, indicate that careful tuning of rate limits is necessary when sending TCP traffic over rate-limited PVCs, as cell drops can have a catastrophic effect on TCP throughput.
- ***IP routing issues*** - Much of the routing complexity was due to the relatively large number of autonomous systems the TPD traffic had to transit. This resulted in a number of special BGP peering arrangements that had to be created specifically to support these demonstrations. Additionally, the parallel satellite/terrestrial paths required special configuration in the U.S. and Japan to ensure a smooth failover to the APAN path in the absence of the satellite link. One of the technical successes of the project was that the complex routing arrangements required were successfully implemented.
- ***TCP window sizes*** - Over a network with a large bandwidth-delay product like a DS3 satellite connection, the amount of data that can be "in-flight" along the link is large—in this case, 3-6 MB. In order to fully utilize this bandwidth, the TCP window sizes must be tunable to these values. Several of the operating systems used in the demonstrations did not support large (>64KB) window sizes as defined in RFC 1323; as a result, the SkyX gateways were necessary in order to improve Visible Human performance results.
- ***NFS/AFS performance tuning*** - There are variables in both AFS and NFS that can be tuned to improve performance and reliability over the underlying infrastructure. Both file systems can be run over TCP or UDP.

Some of these issues were anticipated at the outset, while others did not surface until the project was underway. While it is not possible to anticipate all potential problems, the performance impediments encountered during TPD and the workarounds implemented should be considered in the initial planning for future projects of similar design. This way, major changes to the network architecture during the implementation phase, such as the rearchitecting of the

GSFC and SMU segments to accommodate the SkyX gateways, can be minimized.

Additionally, it would have been useful to have more test equipment in place during the troubleshooting of the layer 2 connectivity along various network segments. For example, only CRL had an ATM analyzer in a location that was useful for ATM testing. As a result, several network problems had to be diagnosed using a combination of hardware and software loopbacks at various ATM switches in the network. The presence of additional test equipment along the segment of the path leading up to the Earth Station would have likely streamlined the troubleshooting process, particularly given the multiple ATM switches that were part of this segment.

4.2 Administrative

Overall, the TPD network illustrated the challenge of coordinating a large number of commercial, government, and academic networks to meet the demonstration requirements. With such a large number of participants, it is important that the requirements be clearly communicated and understood by all parties in order to minimize delays in implementation. In particular, the requirement for dual PVCs over the ATM infrastructure between the BC GigaPOP and CRL/SMU was not clearly communicated.

The TPD network evolved over time from its original design. As a result, new participants continued to join the effort throughout the lifetime of the project. Many of these participants, while giving their best efforts to support the project, could have benefited from earlier involvement in the effort. Again, creating more viable initial designs for future projects (based in part on the technical lessons learned from TPD) should result in earlier identification of and contact with the organizations whose assistance is needed for successful project completion.

5 CONCLUSIONS

TPD illustrated some of the key issues relating to coordination of layer 2 and layer 3 connectivity among multiple international networks, as well as issues relating to the impact of the platform choice on overall performance. Although some of the technical objectives of TPD were not met, the experience gained in the design and implementation of the TPD network will prove beneficial to future projects of this nature.

6 ACKNOWLEDGEMENTS

NREN would like to thank the following for their support of network setup and engineering tests:

- ***GSFC*** – Paul Lang, Kevin Kranacs, Bill Fink, Pat Gary
- ***NLM*** – Mike Gill, Jerry Moran, Jules Aronson, Ivette Reategui
- ***JPL*** – Eddie Hsu, Chaw-Kwei Hung, Claudia de Luna
- ***CRL*** – Naoto Kadowaki, Nozomu Nishinaga, Naoko Yoshimura

- ***UMD*** – Patrick Shopbell, Dan Magorian
- ***SMU*** – Haruyuki Tatsumi, Hiroki Nogawa
- ***MWO*** – Steve Golden, Gil Clark
- ***BC GigaPOP*** – Don McWilliam, Mike Bluett
- ***Teleglobe*** – Normand Forget, Soster Souffrant, George Attia
- ***CA*net3*** – Rene Hatem, Thomas Tam
- ***NISN/CSOC*** – Michael Whisenant, Terry Leger
- ***GRC*** – Will Ivancic
- ***IMnet*** – Mitsutoshi Wada
- ***STAR TAP*** – Doug Pearson
- ***Intelsat*** – Joe Jankowski

USE OF SKYX GATEWAYS IN THE TRANS-PACIFIC DEMONSTRATIONS

Paul Lang

1 LONG-DELAY TUNING AND TESTING, AND USE OF SKYX GATEWAYS

1.1 Rationale for Using the SkyX Gateway

One portion of the Visible Human Demonstration was to demonstrate the use of the Visible Human Slice Server (VHSS). The VHSS was developed on an Apple Macintosh G3 running the MacOS X operating system and uses TCP sockets to transfer the newly generated slice data. Early tests using ftp and nttcp to transfer various size data sets over a simulated two-hop satellite delayed path without being able to use large TCP windows resulted in transfers in the 200 kb/s range or less.

The Mentat SkyX Gateway box was needed to help resolve the problem of poor TCP performance of the VHSS over a long-delay path. The two-satellite hop path results in over a second round trip time delay, and the Operating System (MacOS X) does not allow for the adjustment of the TCP window size above 64K bytes to help minimize the impact of the delay. Previous successful experience with the SkyX Gateway product led to its use in this demonstration.

1.2 SkyX-Related Routing Issues

The SkyX Gateway has two Fast Ethernet interfaces. On the Sapporo Medical University (SMU) end, the SkyX Gateway was placed between the router that connected to the satellite modem and the network on which the SMU test systems were located. The satellite modem in Canada was connected to an ATM switch and therefore did not offer a place for the SkyX Gateway to connect in. Since it was desired to allow SkyX processing for both the National Library of Medicine (NLM) and Goddard Space Flight Center (GSFC) hosts, the SkyX Gateway was placed on a Fast Ethernet at GSFC, and specific static routes were used to force the connections between the demonstrations systems through the SkyX Gateway.

Due to the desire not to add anything between the router and the network to which it is connected, the GSFC SkyX Gateway was set up with an optional SkyX configuration that used only one of its Fast Ethernet interfaces. It was also set up as a one-arm router. Tests were run to verify that the SkyX Gateway could transfer near line rates (96 Mb/s without SkyX processing). Considering that the path was already limited to less than 45 Mb/s, this configuration was determined to be sufficient for this demonstration.

One other routing concern was the importance of having the routes for the SkyX Gateway advertised to go across the two-satellite hop link so that the SkyX Gateways could establish connections between themselves for handling the SkyX processing.

1.3 SkyX Gateway Tuning Issues

The SkyX Gateways are easy to configure. The main information needed is the bandwidth in each direction (rate command) and the delay in each direction (the delay command). Then simple adjustments are made to the rate command to tune the gateway.

Due to configuration issues that were not fully communicated to the group about the rate limiting on the two PVCs, the initial rate configuration settings of the SkyX Gateway were off by a large amount (which had been set to 45000000) and produced less than expected results. With the much appreciated help from Mentat and an undocumented command to limit the bandwidth on the SkyX Gateway, we were able to determine that the effective bandwidth that we were getting was in the range of 16 -17 Mb/s and that the optimal setting of the rate command in the SkyX Gateway configuration should be 16000000 (16 Mb/s) on transmit. This information helped us discover that when a second PVC was added, a 45-Mb/s PVC had been converted into two 20-Mb/s PVCs.

1.4 Benchmark Script

A script was written to help check and save information on the characteristics of the link prior to each Visible Human Viewer test run. The script tested the round trip time, router hops, and transfer rates for a 7-MB file equivalent to the size of the largest slice to be retrieved from the Visible Human Slice Server. The script included the execution of ping (default and 1500 byte packets), traceroute (default and 1500 byte packets from both ends), ftp, and nttcp commands.

1.5 Simulated Satellite

Path	Via	SkyX Process	RTT (ms) 65N/ 1500B	#Hops -> <-	ftp (Mb/s) 15KB/ 32KB	Nttcp (Mb/s) 7MB
ARC-NLM	Simulated	No	1171/1172	2/2	.027/.155	0.203

1.6 Actual Testing:

Path	Via	SkyX Process	RTT (ms) 65N/ 1500B	#Hops -> <-	ftp (Mb/s) 15KB/ 7MB	Nttcp (Mb/s) 7MB
SMU-GSFC	Intelsat	Yes	1124/1127	14/14	/15.2	11.9
SMU-NLM	Intelsat	Yes	1127/1130	16/16	10.9/15.2	11.9
SMU-NLM	Intelsat	No	1127/1130	16/16	.026/.224	0.225
SMU-GSFC	TransPAC	No	191/224	16/14	/.817	0.732

1.7 Troubleshooting

Troubleshooting and coordination in general was complicated—but amazingly overcome—given the differences in time zones, language, additional work loads, holidays, the number of networks/groups/organizations and the mixture of technologies involved.

The use of the large packet traceroute and the archiving of the results helped isolate when and where a burst rate problem was introduced that affected the transmission of large (around 1300 byte or larger) IP packets. The traceroute can send out a large packet but receives a small host unreachable packet. Traceroutes using the default packet size worked in both directions. However, the large packet traceroute worked in one direction but only worked part way in the reverse direction. Large packet pings through the problem link failed in both directions as the entire packet is echoed back.

The use of traceroutes also helped determine when a loop-back had been left in place; the last router before the loop would be repeated in the traceroute.

2 PERSONAL FILE SYSTEM (PFS)

2.1 Rationale for Using PFS

Another portion of the Visible Human demonstration involved retrieving existing slices from a server using a network file system (NFS). NFS was originally tried over the two-satellite link with rather poor results (300 kb/s or less) even with SkyX processing and attempting to change the NFS parameters under MacOS X. As explained by an engineer from Mentat, "SkyX is designed to remove the performance bottleneck imposed by the transport layer (in this case TCP). SkyX cannot do anything about applications that have windowing mechanisms of their own." The PFS test resulted in throughput rates of 2.26 to 8.4 Mb/s, which are up to 28 times better than similar NFS tests.

2.2 NFS Mount by VHP Application (times for 7471284 byte file)

	7MB NFS download in kb/s minimum/average/maximum	7MB NFS download in seconds minimum/average/maximum
Terrestrial	157.71/468.77/591.78	157.71/468.77/591.78
Satellites	143.68/214.19/291.56	205/238/416

2.3 Image Download with PFS (times for 7471284 byte file)

	7MB PFS download in kb/s minimum/average/maximum	7MB PFS download in seconds minimum/average/maximum
Terrestrial	786.70/ 884.64/ 933.24	64.046/67.854/75.976
Satellites	1928.07/5368.55/8413.61	7.104/11.553/31.000

A quote from the PFS homepage (http://www.spa.is.uec.ac.jp/%7Etate/pfs/): "PFS (Personal File System) is a portable network file sharing system designed for mobile computers. It is constructed from file servers on stationary hosts and mobile clients. It has a cache storage on the client, and dynamically adapts for variety of network speeds and bandwidths includes disconnection. All of PFS system is implemented on user land on UNIX, and communicates with client kernel with traditional NFS. Then, PFS can run on variety of UNIX variants."

The detailed results for the PFS testing can be found at: http://www.sapmed.ac.jp/~hakashi/0708results.html

COMMUNICATION RESEARCH LABORATORY—TRANS-PACIFIC DEMONSTRATIONS LESSONS LEARNED

Naoto Kadowaki, CRL

1 NETWORKING INFRASTRUCTURE IN JAPAN—SYSTEM OVERVIEW

Networking infrastructure in Japan is shown in figure 1 (see ***http://www.nren.nasa.gov/tpd)***. The Intelsat Earth Station is located in the Communications Research Laboratory's (CRL) Kashima Space Research Center (KSRC). The Earth Station is operated in Ku-band and directed to transponder 62/22 of the IS-802 at 174 E. The DS-3 modem SDM-9000 was installed to transmit at 44.736 Mb/s data rate with 8PSK modulation scheme. Three ATM switches were located in KSRC. A Fore ASX-200BX is directly connected to the SDM-9000 for the Intelsat link, another is on the line to the CRL headquarters, and the other is connected to the N-Star Earth Station.

KSRC and CRL Headquarters (HQ) are connected by a terrestrial fiber optic link of OC-48 (2.488 Gb/s). The remote astronomy demonstration systems were installed at CRL HQ.

Sapporo Medical University (SMU), which is the main driver of Visible Human demonstration in Japan, is connected to the KSRC by JSAT's N-Star satellite. (Ownership of Ka/Ku band transponders of N-Star-a was transferred from NTT to JSAT in April 2000.) The Ka-band multibeam transponder of the N-Star has a bandwidth of 200 MHz; therefore we could use an OC-3 modem SDM-155 as well as a DS-3 modem SDM-9000 for the connection between SMU and KSRC. We mainly used SDM-9000 for the demonstration because the DS-3 connection is more reliable, but we used SDM-155 for technical checkout or performance tests carried out domestically.

2 INTERNATIONAL CONNECTIVITY

2.1 Technical Review

2.1.1 Bandwidth

The allocated bandwidth on the Intelsat link was 27.3 MHz. The modem that CRL owns did not meet this specification because it employs QPSK R7/8 or QPSK R3/4, and these modulation schemes require more than 30 MHz bandwidth for DS-3 transmission. Therefore, Teleglobe loaned us an SDM-9000 modem with 8PSK modulation.

The transmission signal from KSRC has a bandwidth of 23.5 MHz in 99% power occupying bandwidth. But the transmitted signal's bandwidth was measured as it exceeded the allocated one by about 2 MHz through the Intelsat's Satellite System Operation Guide (SSOG) test. It means that the signal's bandwidth became spread by transponder's nonlinearity. This effect appeared very frequently in satellite communications. But CRL's transmission does not affect any other customers because the transponder had a lot of unused bandwidth in Japan-to-Canada direction. In the Canada-to-Japan direction, the transponder was used almost fully, but interference was avoided by shifting the carrier frequency 3 MHz higher.

2.1.2 Satellite Link Quality

Once we established the satellite link at DS-3 rate, the satellite link was very stable and we did not suffer from bit errors except on July 7 and 8 when a typhoon hit the KSRC area.

2.1.3 ATM Connectivity

ATM testers were set up at KSRC, CRL HQ and SMU. For the first two weeks, we could not receive any cells from North America. After the ATM connection was established, the connection went down, but it was an operational issue, not a technical issue.

2.1.4 IP Connectivity

We tried to use a BGP-4 routing control scheme in order to switch the connection between the terrestrial route and via-satellite route. BGP-4 was applied to the routers at CRL HQ, SMU and BCnet. We could not connect to North America in first three weeks, but after that, the IP connection was established. Once we had the IP connection, BGP-4 worked very well.

2.2 Management review

The Intelsat's SSOG test was carried out by the contract with KDD (Kokusai Denshin Denwa Company).

The Intelsat Earth Station was operated by CRL's researchers in KSRC. Therefore, 7 day-24 hour operation was not easily accomplished. In the first 6 weeks, 24-hour operation was done basically from Tuesday to Saturday. In the rest of the weeks, the Earth Station was operated from 9:00 JST to 21:00 JST on Tuesday to Friday, except for national holidays in relating countries. The satellite link was shut down during the period that the Earth Station was not operated, and this sometimes led to confusion in the operation of terrestrial networks.

We usually operate Earth Stations for experimental purposes, but the operation of the Intelsat Earth Station was a quite different experience. Intelsat is used by many major communications and broadcasting providers, and operation is strictly controlled by the Intelsat Operation Center (IOC). Because the quality of the Intelsat link was controlled under the IOC, we could be notified of abnormal transmission power, bandwidth and other parameters. As the result, we could achieve error free connectivity.

3 DOMESTIC CONNECTIVITY

3.1 Technical Review

3.1.1 Connectivity between KSRC and SMU

JSAT owned N-Star's Ka-band multibeam transponder that was used to connect KSRC and SMU. The transponder has the bandwidth of 200 MHz, and we can obtain OC-3 or capacity with that transponder. CRL installed the Ka-band transportable Earth Station at SMU in January 2000 and tested the link quality in February and March. Cell error performance, TTCP performance and ftp performance are shown in fig.2, fig.3 and fig.4, respectively (see ***http://www.nren.nasa.gov/tpd)***.

3.1.2 Connectivity between KSRC and CRL-HQ

An optical link was used to link KSRC and CRL-HQ, and the capacity allocated to this project was 155.52 Mb/s. This link was stable and error free during the demonstration period.

3.2 Management review

3.2.1 N-Star Linkage

CRL and JSAT exchanged agreement for the usage of N-Star's transponder for this demonstration, and expanded the period as the demonstration period expanded. The Earth Stations in both KSRC and SMU were operated by CRL's engineers.

Operating time was coordinated with tests and demonstrations of the Visible Human demonstration. Originally the N-Star link was operated from 19:00 JST to 10:00 JST, but this period is difficult for CRL's researchers. We shifted the period from 9:00 JST to 23:00 JST after July 12. This was from the same situation as the operation of Intelsat link.

Link quality management was not an issue.

3.2.2 Optical Fiber Linkage

We shared the OC-48 capacity of the optical fiber link between KSRC and CRL-HQ with other CVRL projects. We could not assign an independent PVC for the demonstration, but we had proprietary OC-3 capacity. Therefore we did not have any difficulty with management of this link. The reason why we could not assign the independent PVC is that it takes a long time to register new VPI/VCI through all repeaters between KSRC and CRL-HQ, which is one of drawbacks of terrestrial facilities lack of flexibility.

4 ATM CONNECTIVITY

On the Japan side, the PVC was implemented prior to the start of the demonstration. Once we obtained the PVC between Japan and North America, we implemented it at the ATM switch at KSRC. PVC management was quite easily done, but we faced no international PVC connection for the first five weeks.

Because the satellite link was established for experimental purposes, the link shut down on the weekend. This fact confused communication carriers on the North American side. For example, we detected that the ATM connectivity was disconnected. This situation happened because the carriers found the physical layer was disconnected and set loop-back at an ATM switch. This is one of the issues that we have to solve for future programs.

5 IP CONNECTIVITY

Routers were set up at SMU and CRL-HQ. After the PVC was established, we experienced several problems in establishing end-to-end connectivity. Most of the problems stemmed from lack of recognition between experimenters and communications carriers. Because IP routers were set up at SMU and CRL-HQ only and KSRC did not have a router, we could not verify the domestic connectivity independently. This point should be improved for future programs. After we obtained end-to-end connectivity, we did not have severe problems domestically.

BGP-4 worked very well to switch routes between the terrestrial connection and satellite connection. This was one of the best achievements that allowed us to verify the operability of global-scale IP connectivity.

6 OTHER ISSUES

The link was established for the experimental purposes, so we operated the Earth Stations by CRL staff. We experienced the difference between the operation and experiments for satellite communications. In a pure operation of satellite communication experiments, we established and shut down the link for the data gatherings. On the other hand, the satellite links should be maintained continuously in application experiments and demonstrations. This kind of difference of operation types depends on the objectives of the programs. We should consider the operation scheme for the satellite links for the future programs considering the type of the experiments.

REMOTE ASTRONOMY—TRANS-PACIFIC DEMONSTRATION

Eddie Hsu, Jet Propulsion Laboratory,
and Naoto Kadowaki, Communications Research Laboratory, Japan

1 REMOTE ASTRONOMY DEMONSTRATION

The Remote Astronomy demonstration, as with the Visible Human demonstration, involved the use of Internet Protocol (IP) technologies over a global scale broadband network. The two demonstrations operated over the same global-scale infrastructure. The Remote Astronomy activity, involving high schools in the United States and Japan and the Mt. Wilson Observatory, demonstrated collaborative observation and distance education at multiple locations around the globe and the use of distributed systems technologies. It was designed to facilitate the participation of the general public (students) in the exciting international activities of using satellite communications in a global information infrastructure. It would also help to examine issues in constructing a next generation global/solar-system wide internetwork involving broadband satellites and would provide an opportunity to apply cutting-edge research results from reliable multicast and distributed systems communities. Moreover, the activities help study and develop new technologies and service models, and can span to include activities in global-scale virtual presence, solar system internetwork, disaster mitigation, and other high data rate, distributed applications.

2 REMOTE ASTRONOMY LESSONS LEARNED

2.1 TECHNICAL

2.1.1 Configuration Tools and Distributed Processing

The demonstration involved three primary pieces of software. In addition to the Remote Astronomy software, there were the videoconferencing system and distributed file system technologies. The Remote Astronomy software was developed by Software Bisque for use by the Mt. Wilson Institute. Videoconferencing tools selected included research IP multicast tools and H.323 commercial off-the-shelf tools. The distributed file system software used was the AFS from Transarc. Transarc is, at time of writing, planning to release AFS as open source software.

The Remote Astronomy software accepts multiple TCP connections so as to allow users to monitor on-going activities at the telescope. One user at a time would be able to control the charge-coupled device (CCD) camera to take images through the telescope. The tracking and pointing of the telescopes was facilitated with the use of a Hubble Guide Star Catalog with 19,000,000 objects. The access can be controlled via assigned user name and password for specific periods. In the demonstration, all use the same user name and password; therefore, after taking pictures, a site would need to relinquish the control of the CCD camera for others to use. The coordination instructions originated from JPL by the lead lecturer.

It was planned that research IP multicast tools be used as one of the first steps in leading to the combined use of high definition technology, QoS, layered encoding, and IP multicasting in the longer term. IP multicast also would permit linking the system into the public Multicast Backbone for global access. For this, research multicast tools were readied and verified in the JPL laboratory, and CRL tested UNIX-based multi-site teleconferencing tool "ISABEL" on SGI IRIX machines; however, because of administrative issues, a commercially available H.323 tool, CuSeeMe Pro, was finally used in the demonstration. The H.323 tool was available for PC and Mac platforms. The tool also required the use of a reflector, which is at the central point of a star topology from which receiving sites were sent a copy of the video/audio/message streams from a sender. This reflector, running on a Unix machine, was located at NASA ARC.

The AFS system was selected for the demonstration because of its stability and the relatively low overhead compared to DFS. It helps provide transparent access to users at various locations and facilitated local image processing. There were two AFS servers, one located at NASA ARC and the other at CRL. The images received from the Mt. Wilson telescope would be sent to the NASA ARC AFS server, which would then be replicated to the CRL server. Accesses from Japan would then be local. A week prior to the start of the demonstration, the JPL-supplied Ultra-2 workstation experienced hard disk problems and some data losses. Therefore, the ARC server became the only server in the infrastructure. This meant that more accesses had become remote accesses, making the effects of satellite delay more apparent. Because the AFS itself does not have functions to overcome long-fat pipe's large bandwidth-delay-product, Japanese participants experienced long waiting time at each access to the AFS server. It means that the mirroring scheme or API for full use of extended TCP is required for applications over long-fat pipe connections.

AFS performance tests were conducted using a standard benchmark, Andrew Benchmark by Carnegie Mellon University. IP multicast performance tests were outlined, but were not carried out because the infrastructure was not deployed.

2.1.2 A More Permanent Infrastructure is Important

During the Phase-1 High Definition Video experiment and the Phase-2 demonstrations, a significant amount of time was spent in setting up the global-scale infrastructure. And after the heavy workload and coordination effort, there were only a few weeks’ time in which the team could conduct demonstrations and performance measurements. This brought the need for a more permanent infrastructure with which the participants could conduct experiments and demonstrations.

A more permanent infrastructure would permit the resources committed by all participants to be focused more on experiments, demonstrations, performance measurement and other research activities rather than on coordination and verification efforts to make sure that there was an end-to-end connection.

2.1.3 Need for Flexibility in Configuring Experimental Networks

The current institutional networks were set up for daily operational needs, but setting up an experimental network (e.g., connections, subnets, etc.) was a special activity for institutional networking. While in most instances experimenters were accommodated, it typically takes a long time and does not offer the flexibility sometimes needed on a fast-moving project involving many other participants. Therefore, an integrated institutional plan for accommodating research and experimental networks would greatly help this type of activity.

The high speed connection into the JPL ATM lab was lost and could not be duplicated at ISDS, where the task's equipment relocated. For this JPL was no longer on the high speed backbone (but was still able to participate at a lower data rate and quality of service relative to sites on the high speed backbone). Fortunately, the major capabilities were duplicated at NREN to help support the Remote Astronomy demonstrations, and the Remote Astronomy team used NASA ARC as the primary hub for the capabilities needed for observations. However, this was an example of how advanced network infrastructure removes boundaries and links people together over long distances, and of how reliable services could be provided in a distributed manner for applications such as disaster mitigation.

2.2 MANAGEMENT

2.2.1 Coordination of Telescope

The use of the telescope was coordinated with the Mt. Wilson Institute (MWI). The users of the telescope would schedule a block of time for observations, and MWI would assign a user name and password valid for the duration. For the Remote Astronomy demonstrations, all participating sites use the same user name and password. While the sites can all observe others' activities, care was taken that only one site would control the CCD camera or move the telescope, and this was coordinated by the lead astronomer at JPL. For the purposes of this experiment, the scheme worked smoothly.

2.2.2 Coordination of Engineering and Operations

Weekly meetings were held for the coordination of engineering and operations activities internal to JPL. And two teleconferences were held with the Trans-Pacific team on a weekly basis. The schedules of the JPL activities were linked to the entire team's schedule and were tracked using a computerized management tool.

The use of e-mails was deemed very important, because the activities spanned multiple time zones and countries. Additional conversations were through e-mail and phone calls.

Staff turnover had been a difficult issue and a good amount of time was spent on training to bring new members up the learning curve. Also, the need to tear down the setup in JPL ATM Lab and then configure it again at ISDS consumed significant time and resources.

On the Japan side, the Institute of Space and Astronomical Science (ISAS) of the Ministry of Education coordinated participants for the Remote Astronomy demonstration. However, ISAS could not attend the demonstrations because their satellite was launched during the demonstration period. Soka High School participated in the demonstrations by introduction from ISAS.

2.3 SCIENCE

2.3.1 Soka High School (HS) and Crossroads High School (HS)

In the demonstrations, the participating sites for Remote Astronomy (JPL/Caltech, CRL/Soka, Mt. Wilson Observatory, Crossroads, and University of Maryland) logged into the videoconference and telescope server simultaneously. Participants were able to discuss with each other over the video conferencing system and observe each other's activities in using the remote telescope.

The students belonging to the astronomical observation club in Soka HS had the experience of watching the sky using Mt. Wilson Observatory's online telescope with a telephone line connection. They could thus compare the environments of Internet-based remote observation with telephone line-based remote observation. In their opinion, the Internet-based environment could serve them a lot of material concurrently. For example, they could compare several images and observe participants' faces in the other sites. That environment could be a great help for better understanding and they enjoyed communications with others.

Mini-lectures were conducted from JPL on the structure of galaxies, types and locations of stars, and age and life of stars. Question and answer sessions were held at the end of the observation runs. Objects observed include Messier 5, 10, 64, 86, 87 and 101; Black Eye Galaxy; Great Clusters in Hercules; Eagle Nebula; various star clusters. Images from the Hubble Space Telescope (HST) were used to compare and contrast the images taken by the students during the observations. Some of the observations included:

- ***The Structure of Galaxies:*** Observe examples of the various types of galaxies with the Telescopes in Education (TIE) telescope. Compare with HST observations of similar galaxies, both nearby and very distant. May also compare with Digitized Sky Survey observations of many galaxies.
- ***The Lives of the Stars:*** Observe a number of nebulae with the TIE telescope. Include HII regions (stellar birthplaces), planetary nebulae (death sites of low-mass stars), and supernova remnants (death sites of high-mass stars). Compare with HST observations of similar targets, both in our Galaxy and in others (e.g., the Magellanic Clouds). May also compare with Digitized Sky Survey observations of many nebulae.

- ***Where are all the Stars?*** Observe a number of Galactic star clusters with the TIE telescope. Use these observations, and archival ones from HST, to account for the ~100 billion stars in the Milky Way. Also observe a smaller number of spiral galaxies with the TIE telescope to aid in understanding the larger picture.

Collaborative observations involving sites in Japan and across the United States were demonstrated as participants controlled and imaged celestial objects in turn.

2.3.2 Telescopes in Education Program (T.I.E.)

The Telescopes in Education program (TIE) maintains and utilizes the remote telescopes at Mt. Wilson Observatory, which brings the opportunity to use a remotely controlled telescope and charge-coupled device (CCD) camera in a real-time, hands-on, interactive environment to students around the world. TIE enables students to increase their knowledge of astronomy, astrophysics, and mathematics; improve their computer literacy; and strengthen their critical thinking skills. TIE is a program sponsored by the National Aeronautics and Space Administration (NASA) and developed through the efforts of numerous volunteers, businesses, and supporting organizations including the JPL of the California Institute of Technology. Through TIE, students have rediscovered and cataloged a variable star and assisted the Pluto Express project at NASA's Jet Propulsion Laboratory to revise the ephemeris (orbital location) for the planet Pluto.

Note: See ***http://www.nren.nasa.gov/tpd for*** photos of the Demonstration at CRL Headquarters and the AFS file access from Japan

VISIBLE HUMAN—TRANS-PACIFIC DEMONSTRATION

Mike Gill, National Library of Medicine

1 VISIBLE HUMAN DEMONSTRATION

The Visible Human (VH) Project has its roots in a 1986 long-range planning effort of the National Library of Medicine (NLM). It foresaw a coming era where NLM's bibliographic and factual database services would be complemented by libraries of digital images, distributed over high speed computer networks and by high capacity physical media. Not surprisingly, it saw an increasing role for electronically represented images in clinical medicine and biomedical research. It encouraged the NLM to consider building and disseminating medical image libraries much the same way it acquires, indexes, and provides access to the biomedical literature. Early in 1989, under the direction of the Board of Regents, an ad hoc planning panel was convened and made the following recommendation: "NLM should undertake a first project building a digital image library of volumetric data representing a complete, normal adult male and female. This Visible Human Project will include digitized photographic images for cryosectioning, digital images derived from computerized tomography and digital magnetic resonance images of cadavers."

2 VISIBLE HUMAN LESSONS LEARNED

2.1 Technical

A faster method for file transfer is needed. Even with the higher bandwidth made available by Intelsat and NTT and the delay compensation system, moving a file took too much time for near real-time operation to occur. It was critical for the Japan end to find someway to move a file faster. Late in the experiment they did. They utilized Personal File System (PFS) software. Other methods may be available.

Another issue was that operation of the experimental software at the NLM end was difficult. The prototype segmentation software was in an early stage of development, no user manuals were available, and while the interface evolved over time, at some points it is still in Japanese. The tool was one specialized for use by an anatomist while at the NLM end, personnel handling operations had an engineering background. Having an anatomist collaborator at the NLM end would have proved useful in retrospect. In lieu of that, a better video conferencing system would have helped by facilitating instruction in operation of the software.

Desktop low-end teleconferencing systems were available at both ends (e.g., Microsoft's Netmeeting) and were used to aid in software operation instruction but were generally inadequate in terms of quality due to low frame rate and poor resolution. The audio portion of the collaboration software was useful but it also suffered from poor quality at times. In order to instruct in the operation of the software, remote operation software would have been useful. One party could have taken over control at the other end in order to demonstrate operation to the other when questions came up, new features were added or bugs were encountered. So better collaboration tools would have been useful.

On top of this, communication between Japan and the USA was complicated by the half-day time difference. Email communications were very critical but the delay in response was at least one day typically so there was little opportunity for spontaneous interaction and its potential benefits. In order to make international calls from the USA to Japan, operator intervention was necessary. This limited any benefits that might occur in a spontaneous fashion for two collaborators. For planned meetings this was not an issue.

Another issue that came up included dealing with a prototype operating system. The first iteration of the software operated using Microsoft NT but eventually was ported to Apple's MAC OS-X server. This was a pre-release version and unfamiliar to personnel at NLM, so there were delays for even simple tasks due to the learning process.

2.2 Management

Management took place using teleconferences and through email. As the demonstration progressed the number of parties grew significantly. The message flow between parties working on some specific part of the problem was very large at times. In order to determine the impact of these communications it was critical to stay on top of the message flow. One could not simply ignore them without first opening them. Better critical path information management systems or tools would have been useful in order to maximize information flow to appropriate or involved personnel. Maintenance of a project Web site by a participant was critical to the success of the demonstration.

Enormous amounts of time were expended by numerous parties to establish a link that operated successfully for under a month of time. The satellite portions were very critical but were transitory. In order to study the effect of long links on collaboration activities, more permanent access to satellite links are needed and a management structure put in place designed to facilitate scientific use.

Another issue that came up was related to licensing. The images used were a subset of the Visible Human dataset. Copies of the dataset are available but require the user to have a license. While NLM and SMU were licensees, NASA was not and so had to become a licensee when they wanted to work directly with the dataset.

2.3 Science

This demonstration was an early and necessary step in exploring issues involving interactive biomedical image segmentation, labeling, classification, and indexing done at a distance using large images, specifically the Visible Human dataset. Its primary focus is a Biomedical Image Collaboratory between Dr. Haruyuki Tatsumi of Sapporo Medical University (SMU) and the

NLM. The VH dataset is an information-rich dataset not existing in private sector datasets because commercial subsets of the VH dataset are often compressed by lossy techniques and hence information reduced. By maintaining a centralized repository, management of the resulting database will be more easily done. Updates would be in one place, ensuring authenticity and reliability. Biomedical image libraries (in number and size) are sure to grow. Currently licensees of the VH dataset number 1000+ worldwide. Due to the size and international importance of the dataset, multilingual labeling of the dataset has been proposed. Therefore various researchers are needed to provide image segmentation and labeling. The first such researcher in Japan will work on a lower extremity subset of the Visible Human dataset. In the future online access to an anatomical segmented human anatomy atlas will be a vital resource for biomedical researchers worldwide.

MANAGEMENT LESSONS LEARNED —TRANS-PACIFIC DEMONSTRATION

Edited by Richard desJardins

1 MANAGEMENT LESSONS LEARNED ON THE JAPANESE SIDE

(Naoto Kadowaki, lead)

Language: Because many issues needed to be considered, email and web were essential for communicating with Japanese language speakers (who have much less trouble reading and writing English than speaking English). Although Japanese language speakers were at a disadvantage, the weekly teleconferences were still considered helpful.

Personnel: Both the small staff and limited budget were spread too thin, especially with continuing checkout problems and extensions of the demonstration period of performance.

Hours of operation: Operations coordination was very difficult for the Japanese due to time zone and the International Data Line (i.e., staff had to work at nights and on Saturdays). Since the Earth Station at the Kashima Space Research Center (KSRC) was operated by CRL's researchers, it was difficult to maintain a 24 x 7 operation. The satellite link was shut down during the times when the Earth Station was not operated, leading to confusion in the operation of terrestrial networks.

N-Star linkage: Operating time was coordinated with tests and demonstrations of the Visible Human demonstration. Times of operation were shifted to make it easier for CRL researchers to operate the link.

Intelsat Earth Station: Because the quality of the Intelsat link was controlled under the Intelsat Operation Center (IOC), operators were notified of abnormal transmission power, bandwidth and other parameters, and as a result, could achieve error free connectivity.

Terrestrial link operations: There was no difficulty managing the optical fiber link between KSRC and CRL since they had proprietary OC-3 capacity. An independent Private Virtual Circuit (PVC) could not be established because of the long lead time required to register new Virtual Path Identifier/Virtual Channel Identifier (VPI/VCI) through all of the repeaters between KSRC and CRL HQ. This lack of flexibility is one of the drawbacks of terrestrial facilities.

2 MANAGEMENT LESSONS LEARNED ON THE NORTH AMERICAN SIDE

(Dick desJardins, US Lead, Don McWilliam, Canadian Lead)

Time difference: The time difference complicated communication. Email communications were very critical, but the delay in response was typically at least one day so there was little opportunity for spontaneous interaction and its potential benefits. In order to make international calls from the USA to Japan, operator intervention was necessary. This limited any benefits that might occur in a spontaneous fashion for two collaborators. For planned meetings this was not an issue.

Message flow: Management took place using teleconferences and through email. As the demonstration progressed the number of parties grew significantly. Better critical path information management systems or tools would have been useful in order to maximize information flow to appropriate or involved personnel. Maintenance of a project Web site by a participant was critical to the success of the demonstration.

Cooperative vs command management: TPD was not a project in the usual sense of having funding under the control of a single project management organization. In the future, there should be:

- "Buy in" from many organizations and management
- Continuing "buy in" for prototyping support from commercial provider
- Dedicated systems engineering and integration support

Permanent access to satellite links: In order to study the effect of long links on collaboration activities, more permanent access to satellite links are needed and a management structure put in place designed to facilitate scientific use.

Licensing: The Visible Human images used were a subset of the VH dataset. Copies of the dataset are available but require the user to have a license. While NLM and SMU were licensees, NASA was not and so had to become one when they wanted to work directly with the dataset.

Coordination of telescope: While the sites can all observe others' activities, care was taken that only one site would control the Charge Coupled Device (CCD) camera or move the telescope, and this was coordinated by the lead astronomer at JPL. For the purposes of this experiment, the scheme worked smoothly.

Coordination of engineering and operations: Weekly meetings were held for the coordination of engineering and operations activities internal to JPL. And two teleconferences were held with the Trans-Pacific team on a weekly basis. The schedules of the JPL activities were linked to the entire team's schedule and were tracked using a computerized management tool. The use of e-mails was deemed very important, because the activities spanned multiple time zones and countries. Additional conversations were through e-mail and phone calls.

Technical: It is very difficult to mix commercial network services and experimental network services. Further, if we had used a single 45-Mbps PVP (Permanent Virtual Path), then BCnet could have managed the PVCs (Permanent Virtual Circuit or Connection)

Management: Staff costs of TPD ended up being very high. There were also other staffing issues.